International Journal of Enhanced Research in Science Technology & Engineering, ISSN: 2319-7463
Vol. 2 Issue 8, August-2013, pp: (87-96), Available online at: www.erpublications.com

A Comparative Study on How Big Data is Scaling Business Intelligence and Analytics

Pavan Sridhar, Neha Dharmaji
MS Ramaiah Institute of Technology, BMS College of Engineering, Bangalore, India

Abstract: The term Big Data is causing a furore all around us spanning from news articles to professional magazines, from tweets to YouTube videos, social media and blog discussions. This term, coined by Roger Magoulas from O'Reilly media in 2005, refers to a wide range of large data sets almost impossible to manage and process using traditional data management tools not only due to their size, but also their complexity. Big Data can be seen in the retail, finance and business where enormous data from stock exchange, banking, online and onsite purchasing data flows through computerized systems every day and are then captured and stored for inventory monitoring, customer behaviour and market behaviour. The upsurge in computational and storage power facilitates the agglomeration, storage and analysis of the Big Data sets. Companies introducing innovative and cutting edge technological solutions to Big Data analytics are increasing. The objective of this paper is to study the emergence of Big data and catalog its role in facilitating Business Intelligence and advanced analytics, where techniques such as predictive analytics, data mining, text analytics, statistics, and natural language processing help to understand the current state of the business and track evolving aspects such as customer behaviour to take productive and persuasive decisions. In addition to the underlying data processing and analytical technologies, Business Intelligence and Analytics includes business-centric practices and methodologies that can be applied to various high-impact applications such as e-commerce, market intelligence, e-government, healthcare, and security.

Keywords: Business intelligence (BI), Business analytics, big data analytics, Hadoop, Cassandra, GigaSpaces In memory Data grid (IMDG), NoSQL DB.

Introduction to Big Data: An Evolution not a Revolution

Data insights form an essential part of the decision making process in today's highly competitive business environment. With the massive growth in available data and ways to manage, companies are spending millions of dollars a year on BI, IT infrastructure, transactional applications, BI tools and Business Analytics. Data driven decision making has totally replaced instinct and even reason. Companies can become data-driven, when they place Business Intelligence, Business Analytics and Big Data at the center of their decision making process. It is a big shift in traditional business policies: instead of basing a decision on instinct or experience, each company must begin to analyze its available data. Every two years, the amount of data in the world doubles, and by 2015, it is estimated that the total data on Earth will amount to 7.9 zeta bytes.

Unstructured data, such as text and images accounts for 90% of this amount **[3]**. From here on, it is highly anticipated that this massive amount of data will be used in business analytics to improve operations and offer innovative services. Moreover Business Analytics has been considered by many to be a function of the Business Intelligence process. In the current trend, as Analytics expand and develop sometimes independent from the BI mainframe, some have begun to argue that Analytics is in fact a stand-alone discipline not just a branch of the BI process. Whether that is true or not, however, it is inconsequential. What most business owners must know is that BI combined with Analytics works better and provides a more efficient decision-making process. Figure 1 shows how big data is extensively used in various application segments.

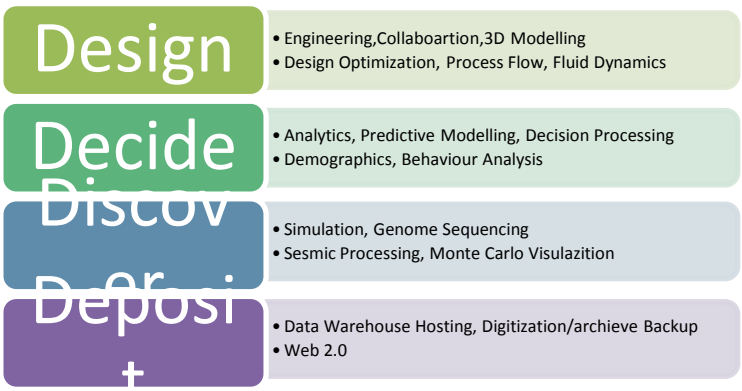


Figure 1: Big Data Application Segment

The four-Vs of Big Data are:

- Volume- Big data comes in one size- large. Enterprises are awash with data, easily amassing terabytes and even petabytes of information.
- Velocity- Often time sensitive big data must be used as it is, to maximize its value to the business.
- Variety- Big Data extends beyond structured data, often including unstructured or semi structured data like text, audio, files images.
- Veracity –Quality and provenance of received data.

How Big Data's Architecture helps Big Analytics

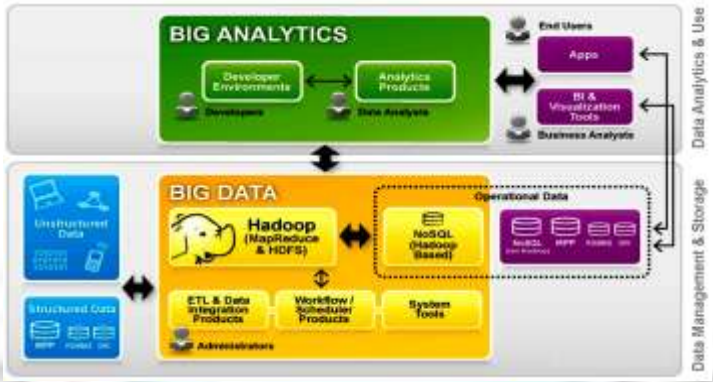


Figure 2: Big Data Architecture

Figure 2 **[28]** depicts a high level architecture of Big Data **[3]**. Let us review well-formed logical information architecture for structured data. Figures 3 & 4 illustrate two data sources that use integration (ELT/ETL/Change Data Capture) techniques to transfer data into a DBMS data warehouse or operational data store, and then offer a wide variety of analytical capabilities to reveal the data. Some of these analytic capabilities include: dashboards, reporting, BI applications, summary and statistical query, semantic interpretations for textual data, and visualization tools for high-density data.

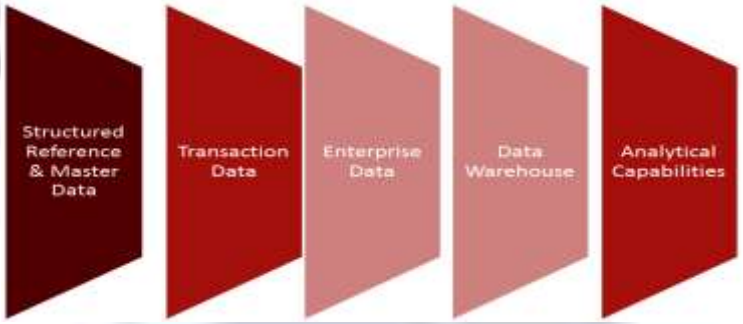


Figure 3: Traditional Information Architectural Capabilities

Unique distributed (multi-node) parallel processing architectures have been created to parse these large data sets **[5]**. There are differing technology strategies for real-time and batch processing requirements. For real-time, key-value data stores, such as NoSQL, allow for high performance, index-based retrieval. For batch processing, a technique known as "Map Reduce," filters data according to a specific data discovery strategy. After the filtered data is discovered, it can be analyzed directly, loaded into other unstructured databases, sent to mobile devices, or merged into traditional data warehousing environment and correlated to structured data.

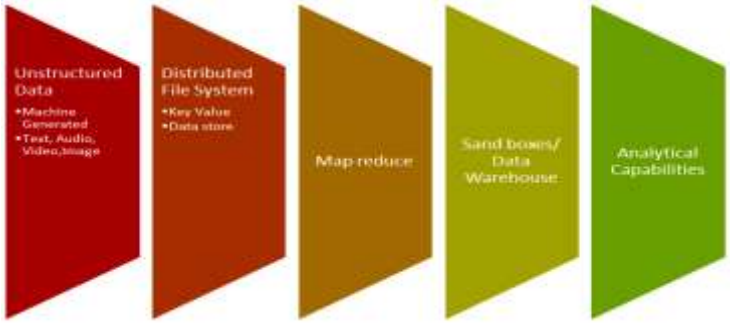


Figure 4: Big Data Information Architecture Capabilities for Unstructured Data

Raw data is not directly moved to a data warehouse. Figure 5 describes Map Reduce processing after which "reduction result" is integrated into the data warehouse environment. It is then leveraged for conventional BI reporting, statistical, semantic, and correlation capabilities. It is ideal to have analytic capabilities that combine a conventional BI platform along with big data visualization and query capabilities. Also, to facilitate analysis in the Hadoop environment, sandbox environments can be created. In summary, the Big Data architecture challenge is to meet the rapid use and data interpretation requirements while at the same time correlating it with other data.

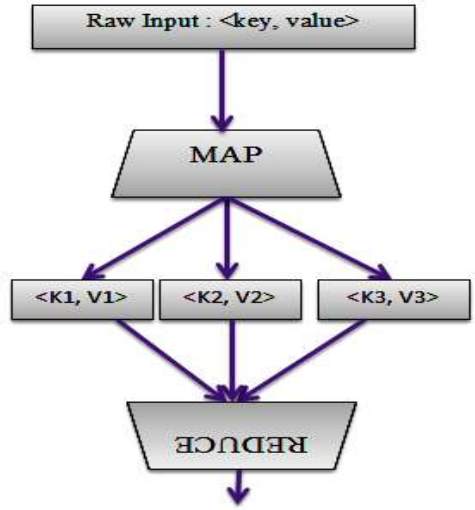


Figure 5: Big Data implementation Flow chart

Technological trends of Business Analytics using Big Data

So far, businesses were limited to utilizing customer and business information contained within an in-house system only. In the past decade however, the role of the Web, e-commerce, social networks, has been significant in tracking users and their preferences. Use of sensors and smart devices are expanding, and more detailed data about people and things are becoming easier to acquire. We are also seeing a rapid increase of individuals disseminating information via social networking services and blogs.The data gathered from such activities lead up to Big Data

With big data, it is necessary to focus not only the volume, but also the variety and velocity of the data. Rather than just single source numerical data, it is necessary to process unstructured data such as text and image acquired from multiple sources. Data that was previously acquired within a number of minutes or hours now use extremely small units of time for acquisition, such as every second, or several hundred milliseconds. Business intelligence (BI) has developed along with visualization in the business environment; however, to utilize big data, visualization is just not enough. Incorporating business analytics (BA) which includes prediction and optimization is the key to success.

There are three major types of business analytics. Type 1 is to find the relationship and regularity between data sets. For example, consumers can be differentiated based on an analysis of the causal relationship between their attributes and purchasing history. Type 2 is to find an optimal solution under a specified set of constraints. This type is valid for problems where limited resources are used effectively, for instance, when optimizing order quantity or scheduling shift workers. Type 3 is to anticipate future trends by understanding guest behaviours **[19]**. In order to realize BA, many IT service providers already offer solutions for large-scale distributed processing (Hadoop), and streaming data processing.

How Big Data Brings BI, Analytics Together

The Business Intelligence and Analytics address the need to make faster and better decisions and increase overall productivity through access to the data and insight required – no matter where the information resides. Analytics helps bring data together, with sophisticated algorithms for filtering and analyzing the data **[4]**. The results can include deep understanding of the workings of the business and its connections to the marketplace, key performance indicators to drive business decisions, and dramatic improvements in the performance of business processes.

"Big Connectivity" plays an integral role in pulling from many different data elements and data sources – seeing further by collecting data and information from processes, applications, Web services, rule sets, social networks, active content, and activities and using all of that Big Data to trigger appropriate changes and actions. It is all about adapting quickly based on as much intelligence and analytics as possible. Figure 6 illustrates the life cycle of Big Data. As the data grows and merges into the cloud, and as the needs to mine that data grow and become increasingly important to business and customer analyses, we see the rise of business intelligence and varieties of data analytics.

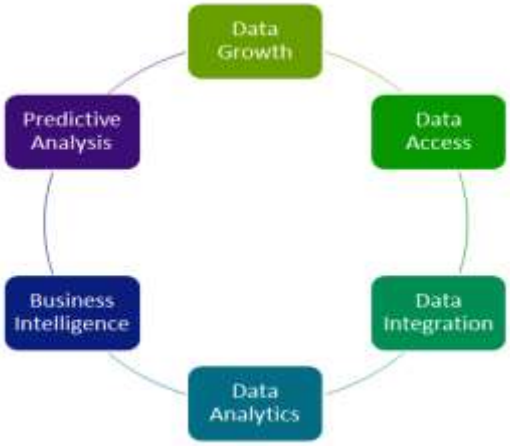


Figure 6: Life Cycle of Big Data

Real time Predictive Analysis Using Big Data

Predictive analytics is mainly comprised of two major components advanced analytics and decision optimization where advanced analytics is grouping of analytic techniques used to predict future outcomes. Advanced analytics includes Predictive Analytics which helps in predicting questions like what will happen next if our customers continue to purchase as they have in the past and how the sales will be impacted if the current trends continues. Hence in this way the analysis is then used to predict future trends, and to spot repeating patterns before they reoccur. That foreknowledge is used to guide business decisions to improve revenue, reduce costs, prevent fraud, and improve customer satisfaction.

Advanced analytics are based on mathematical models and algorithms and started as descriptive statistics which are basically used to sum and count past occurrences for what has happened in the past which is useful in a reactive, course correction manner. Advanced analytics allows you to anticipate possible future outcomes and either capitalize on them or adjust now to impact the future.

The traditional technique for building a predictive model is based on hypothesis testing which more of a statistical approach. Data mining is a technique for building predictive models where the data is visually explored and used to determine which predictive model to use to "fit" the data. For example, if the data visually looks linear then a linear regression technique could be applied. However, if the data plots out logarithmically then a logistic regression technique could be applied. Figure 7 represents how big data is used in various applications of Predictive analytics and the algorithms they use to provide real time business solutions.

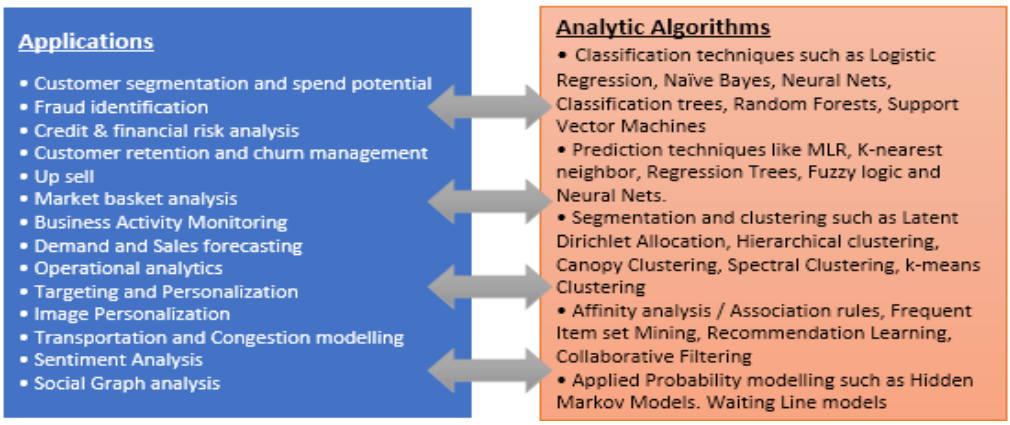


Figure 7: Specialized Predictive Analytics

Precise models are developed using predictive modeling, which help businesses make informed decisions. They are used to:

- Predict competitive initiatives with accuracy
- Identify latent market demand and innovate products to meet it.
- Identify a person's likelihood to develop heart conditions (sentimental analysis) and suggest the right incentives to change contributing behavior.
- Gmail by Google prioritizes mails for its users and groups them into 'Priority Inbox' using predictive algorithms.

Organizations use predictive analytics in every business area to help achieve cost effective, top line revenue growth that translates into real market value for the company.

BI & Big Data Analytics Similarities & Differences

BI is not a new concept. Data warehouses, data mining, and database technologies have existed in various forms for years. Big data as a term might be new, but many IT professionals have worked with large amounts of data in various industries for years. However, now big data is not just about large amounts of data. Digging and analysing semi-structured and unstructured data is new. Fifteen years ago, we did not analyze email messages, PDF files, or videos. The Internet was just a fad; distributed computing was not created yesterday, but being able to distribute and scale out a system in a flash—and within smaller budgets—is new. Similarly, wanting to predict the future is not a new concept, but being able to access and store all the data that is created is new. Various sources claim that 90 percent **[20]** of the data that exists today is only two years old. And that data is growing fast.

Many enterprises have multiple databases and multiple database vendors, with terabytes or even petabytes of data. Some of these systems accumulated data over 30 or 40 years. Many enterprises built entire data warehouse and analytic platforms off this old data. Large retail corporations, such as Wal-Mart, became billion-dollar companies long before big data. So, it wasn't data that drove their business. Data as a service can drive a business- for example Amazon **[6]**. It was an online e-commerce product company. It has now evolved to become a platform as a service, software as a service, big data as a service, and cloud data centre company. Amazon built an incredible recommendation engine over the years from various open source technologies.

Zynga, the Facebook gaming company that is known for hits like Farmville, used Amazon's cloud services to scale its own databases and analytics **[7]**. For data to be useful to users, it must integrate customers with finance and sales data, with product data, with marketing data, with social media, with demographic data, with competitors' data, and more.

Big-Data Real-Time-Performance Analysis in BI&A using Open source tools

Organizations must be well prepared and in a situation to quickly react to the new opportunities and technological challenges that arise. An effective organization today must be able to gather business critical information out of incoming raw data and have it available at fingertips of the decision makes. This process ensures that organisation keeps running and stays competitive. Nowadays companies are dealing with enormous amounts of raw data coming from various sources. The various analytical tools available in market make take minutes, hours, or even days to get the information extracted from the raw data. It is very important to provide the right information in the right context to the right location at the right time in order to give an organization the insight they need to achieve real business agility. Moreover companies are just not focusing on performing analytics anymore; real time analytics is needed.

So in order to perform real time analytics we need to discover new ways like how we use NoSQL DB to perform analytics on big data we need to use GigaSpaces to perform Real-time analytics. GigaSpaces **[12]** a cloud platform allows you to combine the In memory Data grid (IMDG) with a NoSQL DB such as Apache Cassandra from DataStax to perform real-time analytics (Figure 10) A live example for a real-time analytics would be processing market data events coming at an incredible speed (few million events/sec) from the different market feeders. So the real time analytical tools should process such data in real time to perform decisions on buy/sell activities. So these data are further fed to the back testing systems to construct better distinction strategies.

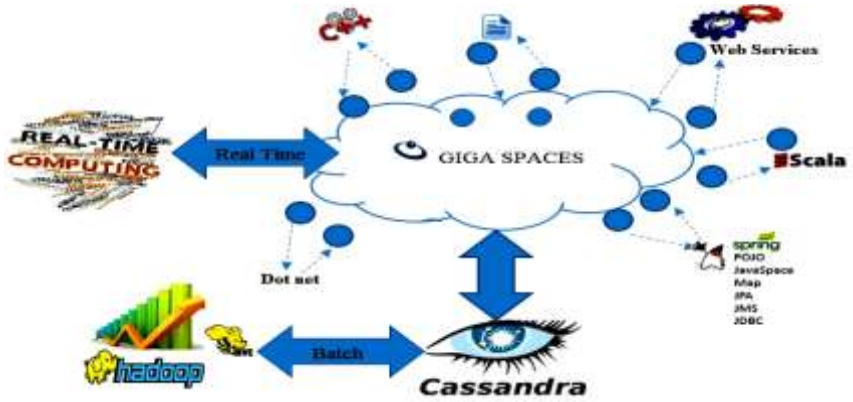


Figure 8 **[12]**: Interaction of GigaSpaces with Cassandra

A two tier architecture is created by combining IMDG and NoSQL where IMDG mainly provides the real time data processing engine where different applications can access the data in real time using different programming languages and software frameworks. The Apache Cassandra NoSQL DB provides the long-term storage for Business Intelligence (BI) use

in real time analytics (via Cassandra) or batch (via Hadoop and Hive). DataStax Enterprise 3.0, a big data platform utilizes a production-ready version of Cassandra for real-time analytics with an integrated Hadoop distribution for batch analysis.

The benchmark results below determines how combining Cassandra NOSQL DB with Giga Spaces IMDG will improve the real time analytics performance for data retrieval operations. The benchmark simply reads data based on a particular key:

Client Threads	Cassandra w/o Giga Spaces TPS*	Cassandra with Giga Spaces TPS
1	3,209	3,398,320
2	5,999	7,156,737
3	8,436	12,302,141
4	11,107	18,368,255
5	13,395	21,476,181
6	15,710	30,230,604
7	17,948	34,364,142
8	21,497	39,376,531
10	20,900	44,381,324

Table 1: Performance chart (* Threads per sec)

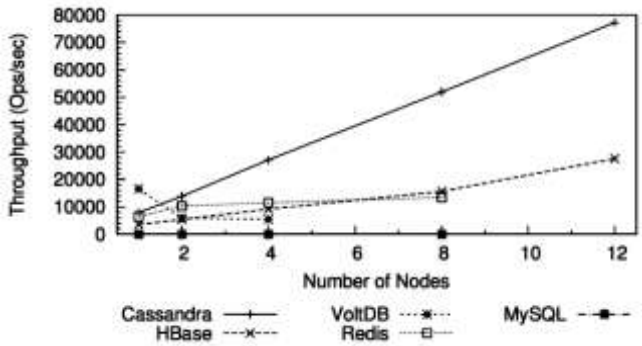


Figure 9: Throughput results for a mixed read, write and sequential scans

In terms of scalability Figure 7 **[31]** clearly confirms us that Cassandra achieves the highest throughput for the maximum number of nodes in all experiments with a linearly steady increasing throughput.

To summarize we can conclude the below points from the benchmark results and graph:

1. Cassandra and GigaSpaces IMDG compatibility has shown to be pronounced, each with different absolute numbers.

2. The advantage of using GigaSpaces IMDG local cache with Cassandra is that it improves the read operation performance by a magnitude of 3(By calculating the slope of the linear line shown in the graph above)

3. Write performance impact on GigaSpaces is very negligible. It takes less than a few microseconds to write the data into the IMDG once it the data is loaded from the No SQL DB.

Case studies of value creation through Big Data Analytics

a) Data Analytics and Visualization using Big Data

Problem Statement: How EBay analyzes huge quantities of data with data visualization providing valuable insight across the company
EBay currently has 108 million active users with $68 billion of goods sold last year; eBay generates a lot of data **[27]**. Employees can access 52 *petabytes* of data on everything from user behaviour to online transactions to customer shipments and much more - with access controls in place to ensure users see only what they're authorized to see. To make all the data understandable, eBay made use of **Tableau [30]**, which provides visualization software to turn large, complex data sets into intuitive, interactive pictures. EBay employees can visualize search relevance and quality or monitor the latest customer feedback and sentiment analysis at any given point of time.

Solution: EBay has built three separate analytic environments, out of which two are data warehouses built on Teradata software with tight security access, and the other is for unstructured data workloads processed using **Hadoop.** Data administered using Hadoop is not made broadly accessible to employees due to the difficulty of creating adequate access

control around it. For structured data such as orders, shipments, listings, bids, payments, customer records, and so on EBay designed a separate 10-petabyte enterprise data warehouse (EDW). The entire organization can connect to the EDW via SQL, and this is where most people pose the bulk of their queries.

EBay have designed a separate data warehouse named "Singularity "as big as 42 petabytes which is mainly used for specialized analytics and storage of structured and semi structured data, which includes the logos and other images from eBay sites. Singularity is mainly used to understand the customer's shopping behaviour on the website and has about 30 percent of the number of EDW's concurrent users.

Results: New advanced analytics tools require specialized skills. Whereas interactive data visualization tools like Tableau can help any business user to become an analyst and identify different kinds of trends on the fly. EBay uses Tableau to visualize search relevance and quality of the eBay.com site; monitor the latest customer feedback and analyse customer sentiments on eBay.com; and achieve operational reporting for the data warehouse systems.
Figure 9 [32] below illustrates how best we can analyse the incoming web traffic data week on week and the page views per visit and thus analyse the daily trends on eBay site. So in this way big data is helping EBay to flourish an analytic culture within the company by making use of business intelligence tools.

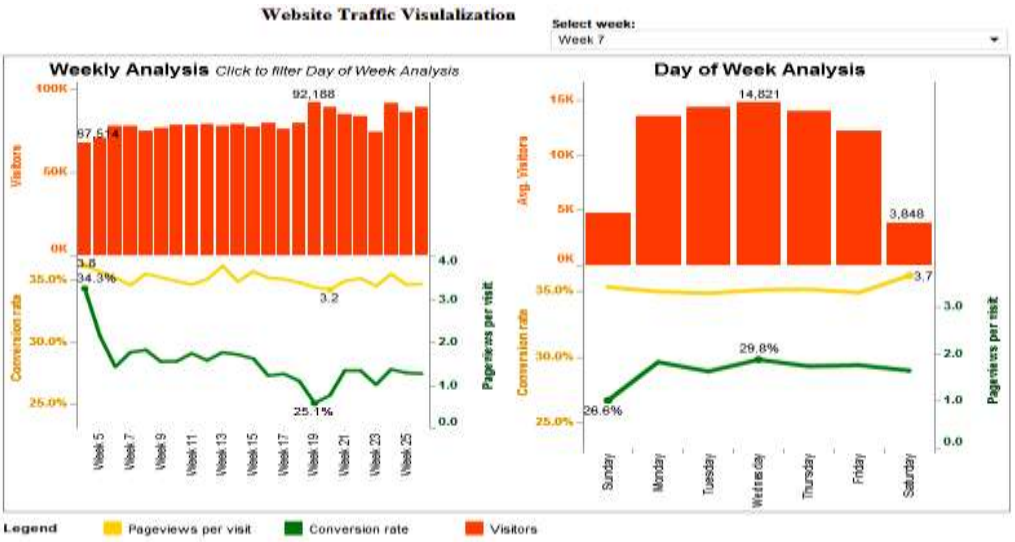


Figure 10: EBay's Web Traffic visualization

b) Web Marketing Analytics Case Study

Problem Statement: Company ABC is a leader in web analytics software and hosted services, with most revenue coming from the latter. It tracks in real-time the activity of visitors to its customer's web sites and provides customizable reports to the customer **[11]**. The company believed that it was missing new opportunities opened up by emerging market trends:

1. Its primary user base has moved from technical web site administrators to non-technical business analysts, while enterprise marketing solutions has become the most lucrative target market.

2. Its customers wanted to track and report on internet properties beyond their corporate web sites. They were also demanding deeper analytical capabilities with faster interactive responses.

3. New competitors were providing basic analytics reporting for free or at low prices, while making money on other businesses enabled by the collected data.

The company changed its business strategy to target business users in the marketing departments of mid to large size enterprises. However it needed to create and implement a new technical strategy to support its new business strategy.

Solution: The company gauged that its technical architecture was outdated and required modernization. When first-deployed, it used a state-or-art proprietary file-based encoding to compactly represent and process web analytics data. It was later extended to provide flexible relational and multi-dimensional reporting using MS SQL Server.

However, the proprietary data formats and engines did not evolve as rapidly as external technologies such as HBase and Hadoop, creating a significant technical disadvantage for the company. The company developed a list of requirements and evaluated several promising data and analytics technologies by implementing proofs of concept prototypes using its own data. Based on this detailed evaluation, it adopted a hybrid architecture based on Cassandra NoSQL database, Hadoop's MapReduce and MS SQL Server. Since the technological landscape was rapidly evolving, the company decides to hedge its bets by shortlisting some attractive alternatives for future reconsideration.

These included proprietary Aster Data System (acquired by Tera Data) and Vertica databases, not selected because of their higher licensing costs; HBase (Hadoop's native database), not selected because of its stronger consistency model and weaker robustness against server failure; MySQL and SQLite, not selected because of their limited support for non-SQL analytical capabilities such as MapReduce. While both Cassandra and Hadoop are open-source, they are also commercially supported by vendors like Cloudera and DataStax.

Benefits: The emerging data platform has already provided several benefits:

1. New analytics capabilities have been released, such as multivariate testing and campaign analytics.

2. The scope of data collection has been expanded to social sites like Facebook and mobile platforms like iPhone. So more scalable than the legacy platform.

3. Software licensing costs have decreased, while the data center operating costs have not grown.

4. Some previously-slow services have been transformed into fast interactive services, such as cross-channel reporting.

Tackling Big Data from Security point of view

Many businesses already use Big Data for marketing and research, yet may not have the fundamentals right – particularly from a security perspective. As with all new technologies, security seems to be an afterthought at best. Big Data breaches will be big too, with the potential for even more serious reputational damage and legal repercussions than at present.

A growing number of companies are using the technology to store and analyze petabytes of data including web logs, click stream data and social media content to gain better insights about their customers and their business. As a result, information classification becomes even more critical; and information ownership must be addressed to facilitate any reasonable classification. Most organizations already struggle with implementing these concepts, making this a significant challenge. Thus data ownership will be distinct from information ownership – perhaps with IT owning the raw data and business units taking responsibility for the outputs. Very few organizations are likely to build a Big Data environment in-house, so cloud and Big Data will be inextricably linked. As many businesses are aware, storing data in the cloud does not remove their responsibility for protecting it - from both a regulatory and a commercial perspective.

Technique such as "**Attribute based Encryption**" is a type of **Public-key** encryption in which the public key of a user and the cipher text are dependent on attributes (e.g. the country he lives, the kind of subscription he has, etc.). In such system, the decryption of a cipher text is possible only if the set of attributes of the user key matches the attributes of the cipher text. So such techniques may be necessary to protect sensitive data and apply access. Many of these concepts are foreign to businesses today.

The security analysis challenge is to find the relevant information in a sea of "noise" – for all the relevant data that can be accumulated; it will usually be dwarfed by vast amounts of unnecessary data. Fortunately, there are specialized techniques that can be used for processing vast amounts of big data to find the useful information. For example, Map Reduce (the programming approach used with data warehouses like Hadoop), whereby problem sets are divided into tasks that can be processed in parallel **[21]**. The interesting opportunity is to see how the application of these techniques to security events and information can improve security insight and proactive management. Big data techniques can help provide context to information, with a view of activities and patterns against particular norms provided in the large data sets.

Results and Summary

Data is very valuable asset of a company. Big Data can help to check and analyze the entire history of an enterprise, interactions with customers' trends and possible threats to the company. Big Data can be thought of as a book with valuable information, and data mining helps to read and understand this "book". Modern BI Solution with Predictive analytics taps this rich vein of experience, mining it to offer something completely different from standard business reporting and sales forecasting: actionable predictions for each customer.
A lot of CFOs and CEOs are leveraging Business Intelligence with its analytic reports in order to have more competitive company. Also, with BI tools companies can be more flexible to risk and fraud and control transactions of the whole chain and one store as well and manage a risk of your business. Traditional business intelligence (BI) extracts relevant data in a structured mart, analyzes it and presents it in formats such as dashboards and reports. Same as data mining, BI tools are more exploratory than action-oriented, but the exploration is more likely driven by a business user than by an analyst. Business Intelligence helps management understand enterprise performance and its trends. While BI tools focus on past performance, predictive analytics forecast behavior and results in order to have specific decisions. Hence, BI shows you what has happened, and predictive analytics tells you what to do. Of course, both are important to making better business decisions [17].

Predictive analytics also work on distilling insight from data, but its main purpose is to explicitly direct individual decisions. Many BI solutions now include some analytics, ranging from report-driven analytics that synthesize past performance data to predictive analytics used in forecasting.

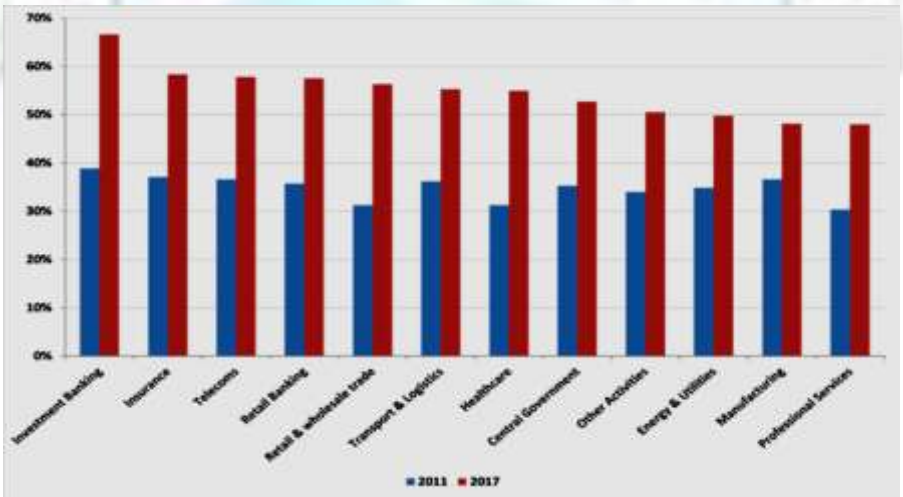
Big data technologies don't have to be complex and require specialized skills. BI tools like SAS provides an extensive array of preconfigured business solutions and business analytics solutions that greatly simplify the most complex analytical problems, including those based on big data. Adoption of more established big data analytics technologies such as advanced data visualization, real-time reporting, advanced and predictive analytics will approach 40-50 per cent by 2015.

Table 2 shows 2011 and Forecast 2012-17 Big Data Analytics Adoption Rates

Year	Big Data Analytics Adoption Rate
2011	34%
2012	38%
2013	45%
2014	50%
2015	52%
2016	54%
2017	54%

Table 2: Forecast 2012-17 Big Data Analytics Adoption

Adoption rates are expected to extend between industries to plus or minus 19 per cent by 2017 from 9 per cent in 2011. It is anticipated that the Financial Services and telecommunication sectors will have the highest rates of adoption by 2017, as these industries are extremely data-intensive and have a higher level of IT spend compared to total turnover. Figure 11 shows current and future rates of big data analytics adoption by industry.



Source: Cebr analysis

Figure 11: 2011 and Forecast 2017 Big Data Analytics Adoption Rates

Conclusion

It is a foregone conclusion that "Big Data" is beginning to take center stage given the explosion of data and the dire need of being able to glean useful business insights from them. "Big Data" Analytics provides the way for identifying useful pearls of wisdom from otherwise useless data. "Big Data" Analytics is becoming mission critical in the enterprises of the future. We have already entered an era of Big Data. Through better analysis of the large volumes of data that are becoming available, there is the potential for making faster advances in many scientific disciplines and improving the profitability and success of many enterprises.

We must support and encourage fundamental research towards addressing these technical challenges if we are to achieve the promised benefits of Big Data. The use of analytics over a wide variety of internal, external and social data (i.e., "big data") allows business processes to become "intelligent," resulting in better business and process outcomes.

So Analyzing new and diverse digital data streams can reveal new sources of economic value, provide fresh insights into guest behavior and identify market trends early on.

But this influx of new data creates challenges for IT departments. To derive real business value from big data, you need the right tools to capture and organize a wide variety of data types from different sources, and to be able to easily analyze it within the context of all your enterprise data.

Acknowledgements

We wish to thank all our colleagues for their constant support and help. We also thank all the managers for their constant guidance.

References

[1]. Hsinchun Chenand, Roger H. L. Chiang, Veda C. Storey A. 2012 "Business Intelligence And Analytics: From Big Data To Big Impact," MIS Quarterly. Rose Business Technologies. Available: http://www.rosebt.com/1/post/2012/11/data-veracity.html June 11, 2012.
[2]. Bernice Purcell. 2012 "The emergence of "big data" technology and Analytics"- Journal of Technology Research.
[3]. Big data: The next frontier for innovation, competition, and productivity. James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela Hung Byers. McKinsey Global Institute. May 2011.
[4]. Data Direct in BI and Analytics Space.2012 "www.datadirect.com"
[5]. Business Analytics Utilizing Big Data – "NTT DATA Technology Foresights 2012"
[6]. Oracle White Paper January 2012: Oracle: Big Data for the enterprise – Jean-Pierre Dijcks
[7]. NTT DATA Technology Foresights 2012 Business Analytics Utilizing Big Data
[8]. Richard L. Villars Carl W. Olofson, Matthew Eastwood - June 2011 "Big Data: What It Is and Why You Should Care".
[9]. Helen Sun, Peter Heller: August 2012 "An Oracle White Paper in Enterprise Architecture: Oracle Information Architecture: An Architect's Guide to Big Data"
[10]. Research Trends "Special Issue on Big Data" Issue 30th September 2012
[11]. Nimble Scale Inc. 2011 "Web Marketing Analytics"
[12]. Big-Data-Real-Time-Performance - Enjoying both worlds with one Architecture! – GigaSpaces Blog - http://blog.gigaspaces.com
[13]. Peter J Jamack 20 Nov 2012 – "Big data business intelligence analytics: Developer Works"
[14]. Philippe Julio – July 2010: "Big Data with No SQL"
[15]. Anton: Business Direction Tools "Predictive Analytics + Big Data = Key Decisions"
[16]. Bitterer, A. 2011. "Hype Cycle for Business Intelligence," Gartner, Inc., Stamford, CT.
[17]. Dr. Dale Skeen: July 2012 "Big Data and Analytics Improve Next-Generation BPM"
[18]. Bernice Purcell- Journal of Technology Research "The emergence of Big Data technology and Analytics".
[19]. Pang, B., and Lee, L. 2008. "Opinion Mining and Sentiment Analysis," Foundations and Trends in Information Retrieval (2:1-2), pp. 1-135.
[20]. Peter J Jamack. 2012 "Big Data business intelligence analytics" IBM developer Works Technical library.
[21]. Perlroth, N., and Rusli, E. M. 2012. "Security Start-Ups Catch Fancy of Investors," New York Times, Technology Section, August 5.
[22]. Russom, P. 2011. "Big Data Analytics," TDWI Best Practices Report, Fourth Quarter.
[23]. Sallam, R. L., Richardson, J., Hagerty, J., and Hostmann, B. 2011. "Magic Quadrant for Business Intelligence Platforms," Gartner Group, Stamford, CT.
[24]. Stonebraker, M., Abadi, D., DeWitt, D. J., Madden, S., Pavlo, A., and Rasin, A. 2012. "MapReduce and Parallel DBMSs: Friends or Foes," Communications of the ACM (53:1), pp. 64-71.
[25]. Michael Cooper and Peter Nell- NIST Information Technology Laboratory, Computer Security Division.
[26]. Sahai, Waters. Eurocrypt 2005. Pirretti, Traynor, McDaniel, Waters. CCS 06. Goyal, Pandey, Sahai, Waters. CCS 06. Kapadia, Tsang, Smith. NDSS 07 – "Attribute Based Encryption".
[27]. Big Data Visualization: A Big Deal for EBay, Available: http://www.infoworld.com/d/big-data/big-data-visualization-big-deal-ebay-208589.
[28]. Big Data, Data Science, Human Capital Analytics, and the Future of Sourcing and recruiting, http://booleanblackbelt.com/2012/04/the-current-and-future-state-of-talent-sourcing2/big_data_sourcing_and_recruiting-2/#.UX0xMLVkPgg April 7, 2012
[29]. Big Data Simplified – Part 3 Available: http://watalon.com/?p=722 March 20, 2012.
[30]. Tableau Software Official Website Available: http://www.tableausoftware.com/learn/stories/ebay-drives-business-analytics.
[31]. DataStax Developer Blog- Available at: http://www.datastax.com/dev/blog/2012-in-review-performance.
[32]. Tableau trail version software downloaded from Tableau official website http://www.tableausoftware.com/solutions/website-analytics.